MOLECULAR AND CELLULAR BIOLOGY, May 1990, p. 1950–1958
0270-7306/90/051950-09$02.00/0

Vol. 10, No. 5

Initiation and Processing of Two Kappa Immunoglobulin Germ Line Transcripts in Mouse B Cells

DEBRA J. MARTIN[1,2] AND BRIAN G. VAN NESS[2]*

Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242,[1] and Department of Biochemistry and Institute of Human Genetics, University of Minnesota, Box 206 UMHC, Minneapolis, Minnesota 55455[2]

Received 18 October 1989/Accepted 10 January 1990

The splicing patterns and sequences of two processed kappa immunoglobulin germ line mRNAs are presented. A 1.1-kilobase (kb) mRNA appeared to be derived from splicing of the previously characterized 8.4-kb germ line transcript, while a 0.8-kb mRNA was the splice product of a second 4.7-kb germ line transcript that initiated 50 base pairs upstream of $J_\kappa 1$. The interaction of the two kappa germ line promoters with nuclear binding factors is also examined. The potential role of these germ line transcripts in establishing the rearrangement potential of the locus is discussed.

Immunoglobulin heavy-chain (IgH) and light-chain (IgL) genes undergo developmentally regulated somatic recombination of multiple gene segments (reviewed in references 2, 3, and 46). These ordered events begin at the IgH locus, with the diversity-gene segment (D) recombining to a joining-gene segment (J), followed by variable-gene segment (V) rearrangement to generate the VDJ_H exon (2, 3, 7, 46, 53). Subsequently, IgL rearrangement of the kappa (Igκ) gene segments V_κ to J_κ occurs. The activation of κ locus recombination appears to require the expression of functional cytoplasmic mu (IgM) protein (2, 3, 53). While it is likely that IgH and IgL rearrangements utilize a common recombinase mechanism, it remains unclear what activates the apparent ordered recombinations of heavy- and then light-chain genes. We (48) and others (52, 53) have suggested that transcriptional activation of germ line (unrearranged) genes may provide an obligatory signal or facilitate rearrangements by opening of local chromatin structure. The accessibility model proposed earlier (52, 53) has been supported by detection of germ line transcripts from murine IgH (16, 28, 35, 52) and IgL (48) germ line gene segments. Analysis of mu-specific cDNA clones from human fetal pre-B cells has shown a significant frequency of sterile (germ line) transcripts (39). Recently, we have also detected germ line kappa transcripts in human B cells (24). A similar correlation between transcription and rearrangement has also been reported relating the induction of murine germ line IgH isotype transcripts to the subsequent switching to the particular isotype transcribed (23, 45). In addition, it has been noted that transcriptional stimulation of artificial constructs enhances the frequency of D to J_H recombination (8, 53).

In Abelson murine leukemia virus (A-MuLV)-transformed pre-B cells (IgM^+) that rearrange the κ locus during cell culture, the unrearranged κ germ line (κ^0) locus is actively transcribed (20). Furthermore, the expression of the germ line C_κ locus can be induced with lipopolysaccharide (LPS) or phorbol ester in many IgM^+ pre-B cells (29). Recently, Schlissel and Baltimore (38) demonstrated that the frequency of kappa rearrangement correlates with the rate of κ^0 locus transcription. Thus, there is strong collective evidence that transcription of the germ line immunoglobulin loci facilitates recombinatorial events.

Interestingly, DJ_H-Cu transcripts yield an mRNA that is translated to yield a truncated IgM protein (35). The functional significance of this IgH protein product is unknown, although it has been suggested that it may play a role in subsequent V_H gene recombination (35). Processed mRNA products originating from the κ^0 transcript(s) have also been observed (17); however, the structure and potential function of these mRNAs are unknown. If processed κ^0 mRNAs also encode protein products, these proteins may play a role in directing κ recombination.

Previous analysis of murine κ germ line transcripts has revealed a transcriptional initiation site that maps to a site approximately 3.5 kilobases (kb) upstream of the unrearranged J_κ cluster (48). In addition, a region that has sequence identity to leader sequences upstream of V_κ coding segments was found 200 base pairs (bp) downstream of this κ^0 initiation site. The presence of a consensus splice site at the 3′ terminus of the leader-encoding region suggested that it may be utilized in subsequent RNA processing. An octamerlike sequence motif common to immunoglobulin gene promoters has also been identified in proximity to the κ^0 promoter (48). Conserved octamer sequences have been found associated with a wide variety of promoters and enhancers (9, 13, 18, 27, 31, 50). For the immunoglobulin genes, it appears to be a major determinant of B-cell-specific expression (9, 13, 18, 27, 31, 50). However, due to the proximity of the κ^0 promoter to the kappa enhancer, transcription of the κ^0 allele may be regulated by other lymphoid-specific factors as well (5, 48).

In accordance with the accessibility model (52), transcriptional activation of the kappa germ line may facilitate opening of the chromatin and subsequent recombination. However, the processed products may contain sequences that would code for translational products that could play a role in the recombination complex. In this study, we report the structure and processing of two murine kappa germ line mRNAs originating from two different nuclear primary RNAs. We examine factors that interact with the promoters located proximal to octamer sequences and discuss a potential role for germ line transcription in B-cell development.

MATERIALS AND METHODS

Cell culture and extract preparation. Cells were grown in RPMI 1640 culture medium (GIBCO) containing 10% fetal

* Corresponding author.

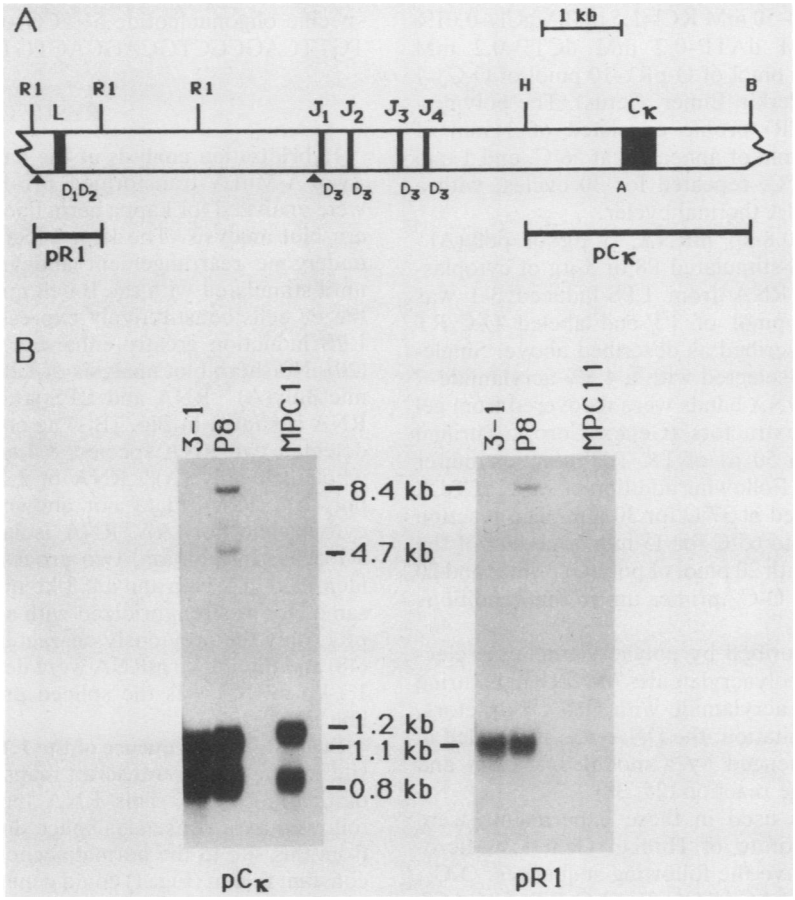

FIG. 1. (A) Restriction map of the κ germ line locus, indicating locations of fragments that were used for specific κ region probes. Restriction enzymes used were *Eco*RI (R1), *Hin*dIII (H), and *Bam*HI (B). D_1, D_2, and D_3 represent different donor splice sites, and A represents the constant-region acceptor splice site. Arrows indicate two sites of transcription initiation described in this report. (B) Expression of κ germ line RNA in murine pre-B cells. LPS-stimulated P8 poly(A)$^+$ RNA (3 μg), LPS-stimulated 3-1 cytoplasmic poly(A)$^+$ RNA (3 μg), or 0.5 μg of MPC11 total RNA was analyzed by Northern blot for κ germ line RNA with the $C_κ$-specific probe $pC_κ$ or the germ line upstream probe pR1. Sizes of RNA were based on mobility of an RNA ladder (BRL) and by comparison with the 1.2-kb and 0.8-kb mRNA products of MPC11 cells (10).

calf serum (Bethesda Research Laboratories [BRL]), 2 mM L-glutamine, 50 μM β-mercaptoethanol, 25 U of penicillin per ml, and 25 μg of streptomycin. A-MuLV-transformed pre-B-cell line P8 was provided by Fred Alt (Columbia University, New York, N.Y.) and has been described previously (36). The 3-1 cells, an A-MuLV-transformed bone marrow pre-B-cell line, was obtained from Robert Perry (Institute for Cancer Research, Philadelphia, Pa.) (29). Pre-B cells were induced by incubation with 15 μg of LPS (Difco) per ml for 16 to 24 h. Nuclear extracts were prepared essentially as described by Dignam et al. (12).

RNA isolation and Northern (RNA) blot analysis. Total RNA was isolated by the procedure of Chomczynski and Sacchi (11). The cytoplasmic fraction was prepared by lysing the cells in 10 mM Tris hydrochloride (Tris-HCl, pH 7.5)–10 mM $MgCl_2$–100 mM KCl–0.2% Triton X-100 and extracting the solution three times with phenol-chloroform (1:1) and once with chloroform, followed by ethanol precipitation. Polyadenylated [poly(A)$^+$] RNA was isolated as described before (6). RNAs were glyoxalated (4), separated on 10 mM sodium phosphate–1% agarose gels, and blotted onto nylon membranes (Nytran; Schleicher & Schuell). Sizes were determined by using RNA markers (BRL).

Probes. The pR1 probe is a 0.9-kb *Eco*RI fragment that lies 3 kb upstream of $J_κ1$ (Fig. 1A) subcloned into pUC19 (48). The $C_κ$ probe $pC_κ$ is a 2.7-kb *Hin*dIII-*Bam*HI fragment cloned into pSP65. The probes were nick-translated with [α-^{32}P]dCTP to a specific activity of 0.8×10^9 to 1×10^9 cpm/μg and annealed with the nylon-bound RNA. The membrane was washed as described before (15) and prepared for autoradiography. The blot was reprobed following removal of the previous probe by washing the membrane three times in 100°C H_2O for 20 min each.

cDNA amplification and sequencing. RNA was purified as described above. For sequencing of the 1.1-kb mRNA, 1 μg of P8 (LPS-induced) total poly(A)$^+$ selected RNA or 1 μg of 3-1 (LPS-induced) cytoplasmic poly(A)$^+$ selected RNA was coprecipitated with 20 pmol of O-$C_κ$ primer. Samples were suspended in 20 μl of 10 mM Tris-HCl (pH 8.3)–250 mM KCl–1 mM EDTA and annealed at 65°C for 2 h. Reaction mixes were made 10 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.3), 5 mM dithiothreitol, 20 U of RNasin (Promega), 0.5 mM dATP, 0.5 mM dCTP, 0.5 mM dTTP, and 0.5 mM dGTP to a final volume of 50 μl. Two hundred units of Moloney murine leukemia virus reverse transcriptase (BRL) was added, and the reaction mix was incubated at 37°C for 1 h. The reaction was stopped by heating to 65°C for 10 min and ethanol precipitated. The pellet was suspended in 100 μl of

10 mM Tris-HCl (pH 8.3)–50 mM KCl–1.5 mM $MgCl_2$–0.01% (wt/vol) gelatin–0.2 mM dATP–0.2 mM dCTP–0.2 mM dTTP–0.2 mM dGTP–20 pmol of O-pR1–10 pmol of O-C_κ–1 U of Taq polymerase (Perkin-Elmer, Cetus). The polymerase chain reaction (PCR) profile consisted of 1 min of denaturation at 94°C, 1 min of annealing at 56°C, and 1 min of polymerization at 72°C, repeated for 30 cycles, with a Perkin-Elmer, Cetus DNA thermal cycler.

For sequencing the 0.8-kb mRNA, 5 μg of poly(A)$^+$ selected RNA from LPS-stimulated P8 or 5 μg of cytoplasmic poly(A)$^+$ selected RNA from LPS-induced 3-1 was coprecipitated with 20 pmol of ^{32}P-end-labeled O-C_κRT primer and reverse transcribed as described above. Single-stranded DNA was size-selected with a 4.5% acrylamide–7 M urea denaturing gel. DNA bands were recovered from gel slices by using Gel/X extractors (Genex Corp.). Purified DNA was suspended in 50 μl of 1× TdT reaction buffer (BRL) plus 1 mM dGTP. Following addition of 9.5 U of TdT, the mixture was incubated at 37°C for 30 min. The reaction was stopped by heating to 65°C for 15 min. One half of the reaction was amplified with 20 pmol of poly(C) primer and 20 pmol of ^{32}P-end-labeled O-C_κ primer under the conditions outlined above.

PCR products were purified by polyacrylamide gel electrophoresis with a 4.5% polyacrylamide–7 M urea denaturing gel and eluted from the acrylamide with Gel/X extractors. Following ethanol precipitation, the DNA was suspended in 50 μl of H_2O and sequenced by a modified Maxam and Gilbert chemical cleavage reaction (25, 38).

Primers. The primers used in these experiments were synthesized by the Institute of Human Genetics Micro-Chemical Facility and have the following sequences: O-C_κ, 5′-ACTGGATGGTGGGAAGATGG-3′; O-C_κRT, 5′-GACT GAGGCACCTCCAGATG-3′; O-pR1, 5′-GGAAAGGACTT GGCTTGTGCCC-3′; poly(C), 5′-GACTCGAGTCGACCCC CCCCCCCCCCCCCCCC-3′; 3′-J_κ1, 5′-AAGAGACTTTGG ATTCTACT-3′. Primers were end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase (BRL).

Primer extension assay. A 2-μl amount of total poly(A)$^+$ RNA from the P8 cells was annealed to ^{32}P-end-labeled 3′-J_κ1 primer and reverse transcribed as described above. A sequencing ladder was generated by using the 3′-J_κ1 primer annealed to pSpECK, a plasmid containing the *Eco*RI-*Bam*HI κ^0 fragment (Fig. 1A) (47), and sequenced according to the specifications given for Sequenase 2 (United States Biochemical Corp.).

Gel retardation assay. The octamer-containing fragments were derived from either a 96-bp *Eco*RI-*Sau*3A fragment isolated from pR1 (see Fig. 4) or a 70-bp *Dde*I-*Sph*I fragment isolated from the pSpECK plasmid. Fragments were ^{32}P-end-labeled with T4 polynucleotide kinase.

The binding reactions (14) were carried out in a total volume of 15 μl containing 10 mM Tris-HCl (pH 7.5), 5% (vol/vol) glycerol, 50 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol, 0.5 μg of sheared salmon sperm DNA, 7 μg of nuclear extract protein, and approximately 5 fmol of ^{32}P-labeled octamer-containing fragment. The assay mixture was incubated for 30 min at 25°C and was loaded onto a 4.5% polyacrylamide gel (acrylamide-bisacrylamide, 25:1) containing 40 mM Tris, 0.384 M glycine, and 2 mM EDTA. The samples were subjected to electrophoresis at 8 V/cm. The gels were then dried and subjected to autoradiography.

Competition assays were performed with the double-stranded octamer-specific oligonucleotide 5′-CCCTGCTG ATTTGCATGTTCCTAG-3′ or the double-stranded non-specific oligonucleotide 5′-TCGACAACCCTAGAGGTTTT TGTTCAGCCCTGCAGCACTGTGGGAGGAACCTCA-3′.

RESULTS

Hybridization analysis of the kappa germ line transcripts. Two A-MuLV-transformed pre-B-cell lines, 3-1 and P8, were analyzed for kappa germ line (κ^0) expression by Northern blot analysis. The kappa loci of the 3-1 cells have not undergone rearrangement and are transcriptionally silent until stimulated with the B-cell mitogen LPS (29). Although the P8 cells constitutively express κ^0 mRNA at low levels, LPS induction greatly enhances κ germ line transcription (20). Northern blot analysis of LPS-stimulated 3-1 cytoplasmic poly(A)$^+$ RNA and LPS-stimulated P8 total poly(A)$^+$ RNA is shown in Fig. 1B. The constant region probe, pC_κ, detected two RNA species, 8.4 and 4.7 kb, from either P8 whole-cell poly (A)$^+$ RNA or LPS-stimulated 3-1 nuclear poly(A)$^+$ RNA (data not shown) that were not seen in cytoplasmic poly(A)$^+$ RNA isolated from LPS-stimulated 3-1 cells. In addition, two processed mRNA species were identified, a 1.1-kb and a 0.8-kb mRNA (Fig. 1B). When the same blot was rehybridized with a probe 5′ of the J_κ cluster, pR1, only the previously characterized 8.4-kb nuclear RNA (48) and the 1.1-kb mRNA were detected, suggesting that the 1.1-kb mRNA was the spliced product of the 8.4-kb germ line RNA (17).

Structure and sequence of the 1.1-kb mRNA. The initiation site for the 8.4-kb transcript maps to a site 3.5 kb upstream of the J_κ cluster. This RNA includes a leader sequence followed by a consensus splice donor site (30, 48). Splicing from this site to the normal acceptor site located before the constant region (Fig. 1) could generate a spliced RNA similar in size to the 1.1-kb RNA. In order to determine the structure of the 1.1-kb RNA, two primers were generated that spanned the potential splice site. One primer, O-C_κ, hybridized to a portion of the coding sequence of the C_κ region, while the other primer, O-pR1, hybridized to sequences located 3′ of the 8.4-kb κ^0 RNA initiation site (yet 5′ of any potential splice sites; Fig. 2A). The O-C_κ primer was end-labeled and annealed to either LPS-stimulated 3-1 or P8 poly(A)$^+$ mRNA, and first-strand cDNA was synthesized by using reverse transcriptase. The O-pR1 primer was added to the reaction mix, and the single-stranded DNA and amplified by PCR. Amplified DNA of the appropriate size for a processed RNA was obtained and subjected to modified Maxam and Gilbert sequencing reactions (26, 42). A similar reaction with unlabeled O-C_κ and end-labeled O-pR1 was performed in order to obtain sequence from the opposite direction. Figure 2 shows the complete sequence obtained for the 1.1-kb mRNA, starting from the previously described initiation site (48) of the 8.4-kb RNA and terminating at the poly(A) addition site for the constant region. Termination codons were found in all three reading frames. An interesting feature of this RNA was the selection of splice donor site. Previous in vivo studies support the scanning model (22, 33), in which splicing occurs from the 5′-most available donor site. Contrary to this model, the 1.1-kb RNA was not spliced from the 5′-most donor site located at the 3′ end of the leader sequence (D_1, Fig. 1 and 2); rather, a splice donor site 137 bp downstream of the leader splice site (D_2, Fig. 1 and 2) was used.

Identification of a second initiation site and sequence analysis of the 0.8-kb mRNA. As shown in Fig. 1B, when total poly(A)$^+$ RNA isolated from P8 was hybridized with the pC_κ probe, a 4.7-kb transcript was detected. The transcript was

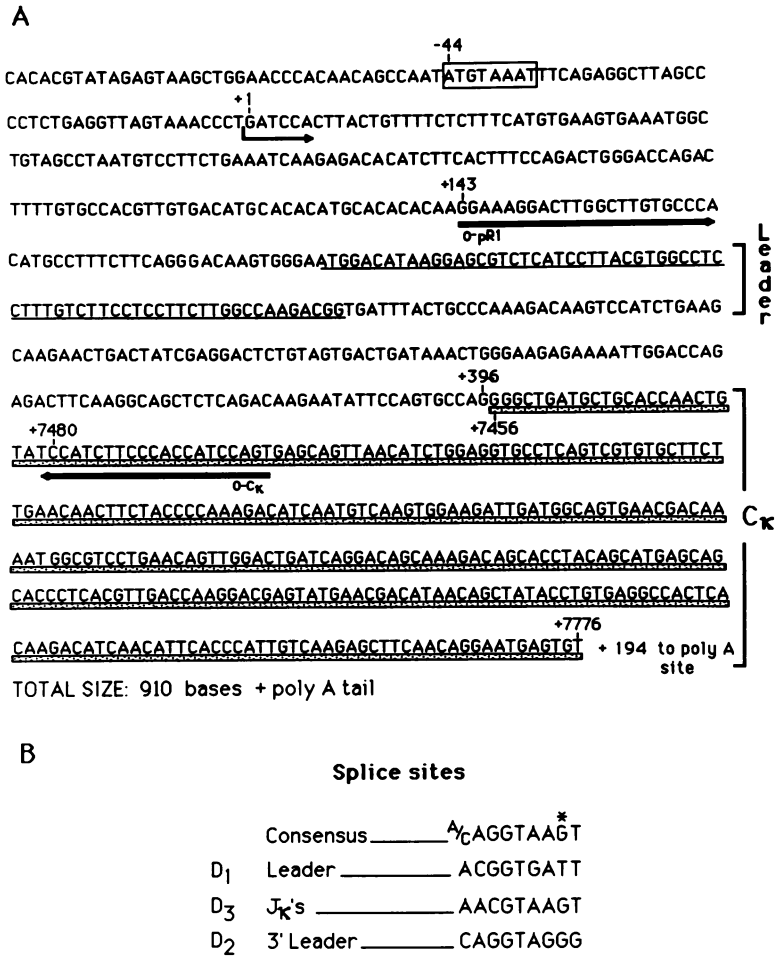

FIG. 2. (A) Nucleotide sequence of the 1.1-kb κ germ line mRNA. Boxed sequence indicates the location of an octamer sequence important in promoter regulation. Primers used for the PCR amplification, O-pR1 and O-C_κ, are boldly underlined, with arrows indicating orientation. Transcription initiation site is represented by a bent arrow. The leader sequence characterized by Van Ness et al. (48) is underlined, and the constant region is underlined with a stippled line. (B) Splice site sequences of the 8.4-kb κ^0 transcript. The asterisk marks the G residue that if mutated causes suppression of proper splicing.

not detected in the LPS-stimulated 3-1 cytoplasmic RNA fraction but could be seen in the LPS-stimulated 3-1 nuclear RNA fraction (data not shown). Since this transcription did not hybridize to the pR1 upstream probe, which spans the 8.4-kb κ^0 initiation site, it is not likely to be a splicing intermediate or product of the 8.4-kb κ^0 RNA. Because the P8 cell line undergoes rearrangement of the κ locus in culture (35, 36), the 4.7-kb RNA could be the precursor RNA from a subpopulation of cells that had rearranged the κ locus. However, when compared with known sizes of $VJ_\kappa 1$-4 primary transcripts, the 4.7-kb RNA fell between the transcripts generated by $V_\kappa J_\kappa 2$ (5.0 kb) and $V_\kappa J_\kappa 3$ (4.4 kb) rearrangements (33). Analysis of the known DNA sequence (25) approximately 4.5 kb upstream of the poly(A) addition site [4.7-kb mature message minus 200 bp of poly(A) tail] identified an octamerlike motif 147 bp 5′ of $J_\kappa 1$ (Fig. 3). The presence of the octamer 5′ of $J_\kappa 1$ suggested that this region could contain a second promoter for κ^0 transcription. A primer extension assay utilizing a primer located between the $J_\kappa 1$ and $J_\kappa 2$ exons identified a precursor RNA with an initiation site located 51 to 53 bp 5′ of $J_\kappa 1$ (Fig. 3A). This result suggests that transcription initiates from the sequence CT*TC*A . . . (where asterisks indicate start site) located approximately 95 bp downstream of the octamer. S1 nuclease mapping with a DNA fragment containing the 80 bp 5′ of $J_\kappa 1$ confirmed this result by protecting a 50-bp fragment (data not shown). These data therefore define a second initiation site for κ^0 transcription 52 bp upstream of $J_\kappa 1$. This location is consistent with the region's being a promoter for the 4.7-kb transcript.

To investigate whether the 0.8-kb mRNA was a processed product of the 4.7-kb κ^0 RNA, we conducted experiments similar to those described for the 1.1-kb κ mRNA. Poly(A)$^+$ RNA from either LPS-induced 3-1 or P8 cells was annealed to end-labeled O-C_κRT primer, and single-stranded cDNA was synthesized. The cDNA was then poly(G) tailed at the 5′ end with TdT and amplified by PCR to generate enough material for sequence analysis (see Materials and Methods). Similar analysis of both cell lines identified an mRNA species that initiated 5′ of the $J_\kappa 1$ recombination signal sequence (heptamer/nonamer) and was spliced $J_\kappa 1$ to C_κ (Fig. 3A). Combining the results of both the primer extension and cDNA sequencing, the complete sequence of the 0.8-kb mRNA was determined and is presented in Fig. 3B. Although this splicing creates an extended open reading frame, no initiation AUG is present for this open reading frame in the 0.8-kb RNA.

Identification of octamer potential. The octamer motif

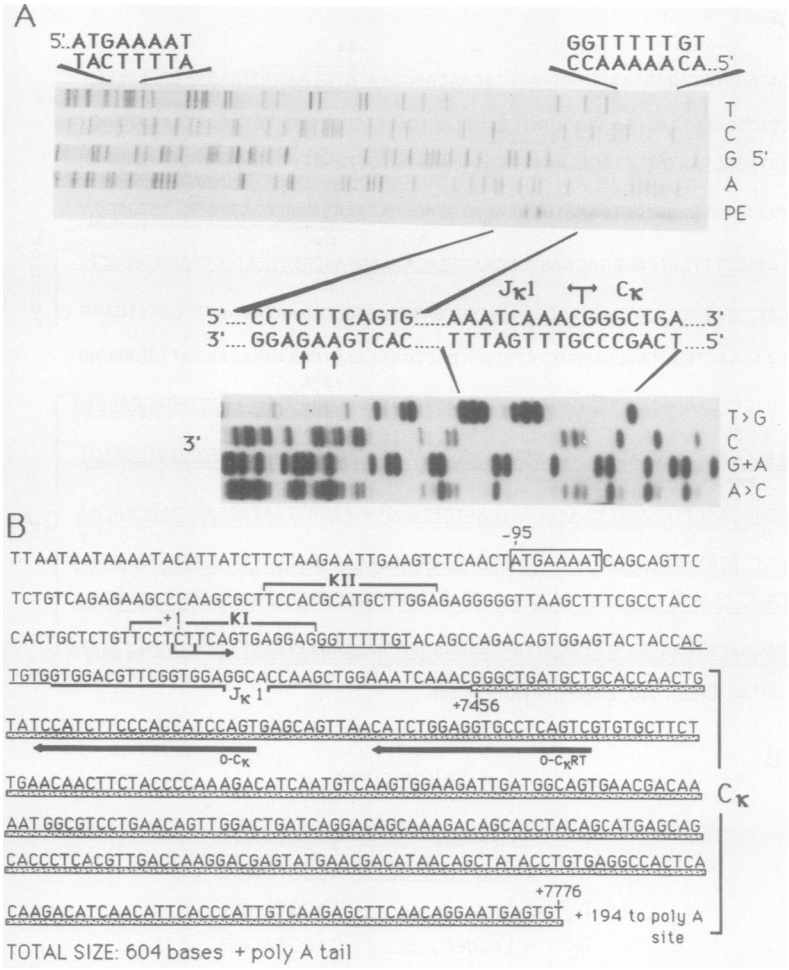

FIG. 3. (A) Primer extension and Maxam-Gilbert sequencing gel of the 4.7-kb κ^0 transcript and the 0.8-kb κ^0 mRNA. A 3′-J_κ1 primer was annealed to 1 μg of P8 total poly(A)$^+$ RNA and extended with reverse transcriptase (BRL). As a marker, the 3′-J_κ1 primer was annealed to pSpECK (45) and sequenced. Sequencing ladder and primer extension (PE) are shown in the upper panel. Locations of an octamer sequence, ATGAAAAT, and the recombination recognition nonamer, GGTTTTTGT, are indicated. Initiation sites identified by primer extension are indicated with arrows. Lower panel is a sequencing gel showing the junction of J_κ1 to C_κ. (B) Nucleotide sequence of the 0.8-kb germ line mRNA. Boxed sequence indicates the location of an octamer sequence. Primers used for reverse transcription (O-C_κRT) and PCR (O-C_κ) are boldly underlined, with arrows indicating orientation. Binding sites KI and KII for the nuclear protein KLP are marked (49). Initiation sites are represented by bent arrows. Heptamer-nonamer recombination recognition sequences are underlined and are adjacent to the more boldly underlined J_κ1 region. The constant region is underlined with a stippled line.

ATGCAAAT has been implicated in the regulation of tissue-specific expression of immunoglobulin genes (9, 13, 18, 27, 31, 50). Van Ness et al. (48) identified an octamerlike sequence, ATGTAAAT, located 44 bp 5′ of the 8.4-kb κ^0 RNA transcriptional initiation site. This sequence differs from the consensus octamer motif at position 4 (underlined; C to T). There has been no direct evidence that this sequence participates in binding of the well-characterized octamer-binding proteins (19, 37, 40, 44). In order to determine whether this sequence is able to bind nuclear proteins, a 96-bp *Eco*RI-*Sau*3A fragment (Fig. 4A) containing this octamerlike sequence was used to investigate whether it bound nuclear protein previously identified for classical octamer sequences. This fragment was end-labeled and subjected to the gel retardation assay with LPS-induced or noninduced 3-1 nuclear extracts (Fig. 4A). The DNA-protein shift pattern is characteristic of the NFA-1/NFA-2 shift seen by Staudt et al. (44), in which the ubiquitous NFA-1 binding is observed in both extracts, while the lymphoid-specific octamer-binding protein NFA-2 is detected only in the LPS-induced extract. The difference in the mobility of NFA-1 when a DNA fragment is compared with an oligonucleotide is probably due to the difference in the sizes of the probes. To verify that these shifts represented binding of NFA-1 and NFA-2, gel retardation assays were conducted with a consensus octamer oligonucleotide as the competitor. As shown in Fig. 4A, binding to both NFA-1 and NFA-2 was competed with by an unlabeled oligonucleotide containing the consensus octamer sequence, but not by a nonspecific oligonucleotide.

The octamer motif ATGAAAAT located 5′ of the 4.7-kb κ^0 RNA also differed from the consensus octamer by a single base. When a 70-bp *Dde*I-*Sph*I fragment containing this octamer was subjected to gel shift analysis as above, a different banding pattern was observed with an inducible factor (B4) that could be competed with by the unlabeled consensus octamer oligonucleotide but not by a nonspecific oligonucleotide (Fig 4B). Interestingly, this DNA-protein

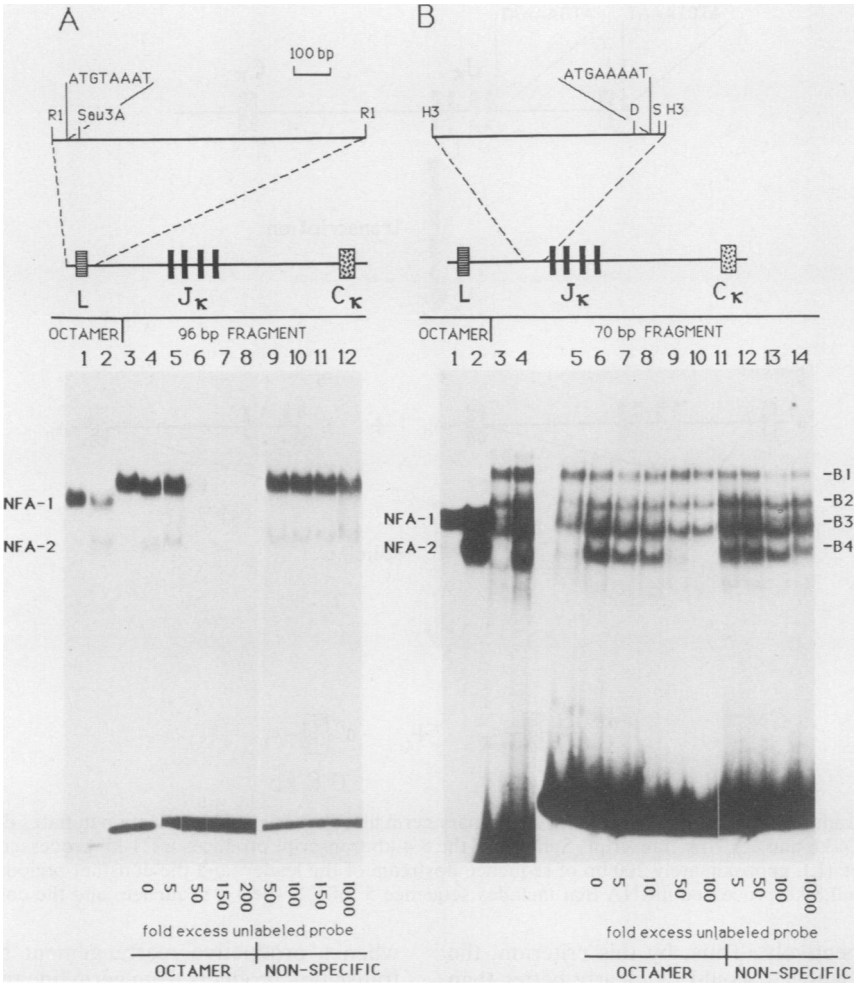

FIG. 4. Gel retardation assays of the germ line promoters. (A) An *Eco*RI (R1)-*Sau*3A 96-bp fragment was excised from pR1 and subjected to gel retardation assays (5 fmol) to detect nuclear proteins with binding specificity for the octamer sequence. Lanes 1 and 3 are nuclear extracts isolated from uninduced 3-1 cells. Lanes 2 and 4 to 12 are nuclear extracts isolated from LPS-induced 3-1 cells. Lanes 1 and 2 show binding to a 24-bp oligonucleotide that contains the consensus octamer sequence. Lanes 3 through 12 show binding to the 96-bp fragment. An unlabeled 24-bp consensus octamer oligonucleotide (5-, 50-, 150-, or 200-fold excess [lanes 5 to 8, respectively]) or a 54-bp nonspecific oligonucleotide (5-, 100-, 150-, or 1,000-fold excess [lanes 8 to 12, respectively]) was used to compete with the 96-bp fragment for nuclear protein binding. (B) An *Sph*I (S)-*Dde*I (D) 70-bp fragment was excised from a *Hin*dIII (H3) fragment isolated from pSpECK and subjected to gel retardation assays (5 fmol). Lanes 1, 3, and 5 are 3-1 uninduced nuclear extracts. Lanes 2, 4, and 6 to 14 are 3-1 induced nuclear extracts. Lanes 1 and 2 show binding to a 24-bp octamer oligonucleotide. Lanes 3 to 14 show binding to the 70-bp fragment. The unlabeled 24-bp consensus octamer oligonucleotide (5-, 10-, 50-, or 100-fold excess [lanes 7 to 10, respectively]) or the 54-bp nonspecific oligonucleotide (5-, 50-, 100-, or 1,000-fold excess [lanes 11 to 14, respectively]) was used to compete with the 70-bp fragment for nuclear protein binding. DNA-protein complexes (B1 through B4) are indicated.

pattern is identical to that reported by Landolfi et al. (19), who used a DNA fragment containing the octamer sequence associated with IgH variable-region genes.

DISCUSSION

The results presented define the structure of two processed kappa germ line mRNAs. Our results are consistent with the 1.1-kb mRNA being derived from splicing of the previously characterized 8.4-kb germ line transcript (48), while the 0.8-kb mRNA is the spliced product of a 4.7-kb nuclear κ RNA that initiates upstream of $J_κ1$ (summarized in Fig. 5). An interesting feature of the 1.1-kb structure was the splice site selection. The $κ^0$ 8.4-kb primary RNA contains six potential splice donor sites (Fig. 1A). Four that are located after each $J_κ$ gene segment (D_3, Fig. 1A) are known to be functional in $V_κJ_κ$-$C_κ$ splicing. Although in vitro studies found no distinction in $J_κ$ splice site usage (22), in vivo studies identified only products arising from a splice to the 5′-most J segment (22, 33). Thus, for the 1.1-kb RNA, it might have been expected that the 5′-most splice site (D_1, Fig. 1A) would be used. Unexpectedly, a 3′ donor splice site (D_2, Fig. 1A) was utilized. When the consensus donor splice site, (C/A)AG:GTAAGT (1, 30) was compared with $J_κ$ donor sites D_3 (AAC:GTAAGT) (25), D_1 (ACG:GTGATT), and D_2 (CAG:GTAGGG) (Fig. 2B), the order of best homology was the $J_κ$ regions (8 of 9), D_2 (7 of 9), and D_1 (6 of 9). The relative strengths of these donor sites could be quantitated by using the score formula outlined by Shapiro and Senapathy (41). According to this formula, the scores for D_3, D_2, and D_1 are

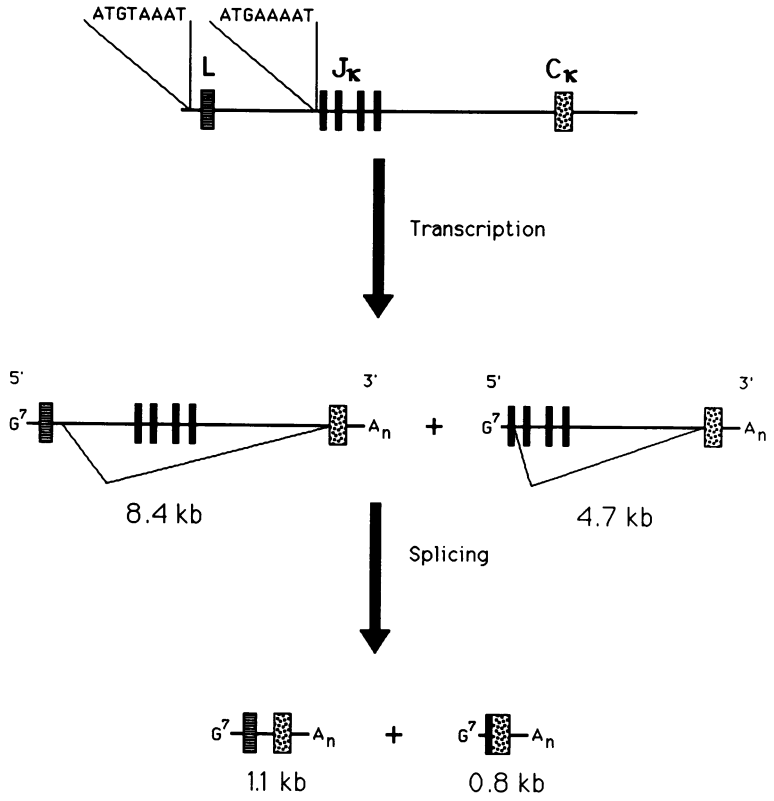

FIG. 5. Summary of transcription and splicing for the two primary germ line transcripts. Transcription initiates downstream of the octamer sequences, yielding an 8.4-kb and a 4.7-kb transcript. Splicing of the 8.4-kb transcript produces a 1.1-kb processed mRNA that contains the germ line leader sequence (L), approximately 200 bp of sequence upstream of the leader, and the constant-region sequence. Splicing of the 4.7-kb transcript yields a 0.8-kb processed mRNA that includes sequence 5′ of $J_\kappa 1$, the $J_\kappa 1$ sequence, and the constant-region sequence.

86, 82, and 75%, respectively. Thus, by this criterion, the sequences associated with D_2 would be clearly better than D_1 and therefore would be preferentially selected. This would be in accord with the splice site strength models (34). Another possible explanation for not observing splicing from the leader site (D_1) comes from work by Zhuang et al. (54), who showed that mutations of the G residue in position 8 of the consensus sequence cause suppression of proper splicing. Notably, the leader splice site does not have a G in the eighth position, implying it could represent a nonfunctional donor site. This would make D_2 the 5′-most functional splice site and would thus be consistent with previous in vivo results (22, 33) and our data.

Using a primer located 3′ of the $J_\kappa 1$ splice site, we confirmed by primer extension that the 0.8-kb mRNA originated from a primary RNA which initiated 5′ of $J_\kappa 1$. Therefore, the 0.8-kb κ^0 mRNA is not an alternative splice product of the 8.4-kb κ^0 transcript. The promoter region for the precursor RNA was suggested by lymphoid-specific nuclear protein interaction (see below).

Sequence data from the 1.1-kb κ^0 mRNA identified no open reading frames, while the 0.8-kb κ^0 mRNA sequence indicated an open reading frame that did not have an initiating AUG codon. Therefore, these mRNAs do not encode a protein product. Interestingly, these results are similar to the observation of Lennon and Perry (21) for the C_μ transcript that initiates within the J_H-C_μ intron. As they suggested, the creation of a nontranslatable exon (nontron) may be important for the mechanism of allelic exclusion (29, 53). An important feature of this mechanism is a feedback system that ceases rearrangement of immunoglobulin genes when a productive rearrangement has been made. Thus, translation products from germ line transcripts could repress rearrangement of immunoglobulin genes prior to the formation of a productive allele. If transcription of germ line alleles is an essential component of the rearrangement mechanism (2, 8, 9, 48, 53), then the production of a nontron would not prevent rearrangement from occurring by sending a premature feedback signal.

A recent report by Leclercq et al. (20) also characterized a 0.8-kb κ mRNA from P8 cells. Based on RNase protection mapping and cDNA library sequence data, the initiation site was mapped to a region 105 to 115 bp 5′ of $J_\kappa 1$. This region includes a sequence, KII, that was found by Weaver and Baltimore (49) to interact with the DNA-binding protein KLP. A similar site, KI, located 5′ of $J_\kappa 1$ also binds KLP (Fig. 3B). Interestingly, using different methods (primer extension and S1 mapping), we mapped the 4.7-kb primary transcript, which spliced to make a 0.8-kb mRNA, initiation site to the KI sequence. It is not clear why the data are conflicting; however, the significance of initiation sites originating from two sequences that bind the same protein suggests a possible transcriptional relationship between KLP and the 4.7-kb κ^0 transcript. Consistent with its proposed role in recombination, KLP was found in pre-B cells and early B cells (49). Currently, the exact function(s) of KLP is unknown, but, from the above data, one could speculate that KLP and κ^0 transcription are intimately involved. Since KI and KII are proximal to the $J_\kappa 1$ heptamer-nonamer recombination signal sequences, KLP may regulate κ gene rearrangement by its involvement in transcription of the 4.7-kb κ^0 RNA and possibly by targeting the

recombinase to this region. This may explain why $J_\kappa 1$ is preferentially seen as the site of V_κ recombination (51).

At least two other distinct DNA sequence elements influence transcription of the κ immunoglobulin gene. One of these is the enhancer located in the J_κ-C_κ intron (see reviews in references 9 and 27). The other essential element is the octamer motif ATTTGCAT, which is found upstream of all V_κ gene transcription initiation sites (9, 27, 31). This element is also present in an inverted orientation at the same location in κ germ line and IgH promoters (9, 27). Recent experiments by Parslow et al. (31) suggest that the octanucleotide may mediate interaction among promoters and enhancers. Unlike the V_κ octamer, which must first rearrange to interact with the enhancer, the κ^0 octamer(s) is already positioned near the enhancer.

The promoter regions for the κ^0 transcripts may be similar to V_κ promoters, based on the presence of the TATA sequence and octamer sequences, yet binding of nuclear factors has not been described. Three proteins are known to recognize the octamer motif, the ubiquitous NFA-1, and the lymphoid-specific proteins NFA-2A and NFA-2B. Interactions of these factors are postulated for B-cell-specific activity of immunoglobulin genes (37, 44). The octamer located in the 8.4-kb κ^0 transcript promoter showed a binding pattern similar to that observed for the V_κ octamer, suggesting NFA-1/NFA-2 involvement. Atchison and Perry, utilizing gel retardation assays, DNase I footprinting, and methylation interference analyses, have found nuclear factor binding specifically at the octamer sequence (M. L. Atchison and R. P. Perry, personal communication). This is interesting since the κ^0 octamer ATGTAAAT has one base different from the V_κ octamer consensus sequence. The octamer of the 4.7-kb κ^0 transcript promoter, ATGAAAAT, is also one base different from the consensus sequence. Although the binding pattern was different from the predicted NFA-1/NFA-2 pattern, it does resemble the pattern seen by Landolfi et al. (19) for the V_H octamer sequence. Moreover, an inducible factor which has the same migration as NFA-2 (Fig. 4B, lanes 2 and 4) was detected and specifically competed with the consensus octamer sequence. Our data strongly suggest that regulation of transcription from the unrearranged κ allele may be similar to that seen for the rearranged allele, although this needs to be directly verified by functional studies.

The results reported in this paper suggest that since both the 1.1-kb and 0.8-kb κ germ line mRNAs lack an open reading frame or AUG start codon, there is not a κ germ line protein product. This is unlike the DJ_H-Cμ transcript, which is translated to produce a protein product (35). Still, the evolutionary maintenance of κ germ line transcription strongly suggests that it plays an important role in B-cell development. Schlissel and Baltimore found that LPS induction of germ line transcription increases the frequency of κ rearrangement (38). LPS is also known to activate binding of nuclear proteins such as NFA-2 and NF-$_\kappa$B (5, 14, 41, 43, 44) to *cis*-acting DNA sequences within the promoter and enhancers. We have demonstrated that the κ germ line promoters bind to LPS-inducible lymphoid-specific nuclear factors. Although the signal(s) for recombination appears to correlate with the transcriptional activation of a specific immunoglobulin locus, it has also been suggested the binding of nuclear factors such as NFA-2, NF-$_\kappa$B, and KLP may contribute to recombination activation (38, 49).

ACKNOWLEDGMENTS

We are grateful to Rod Feddersen for his suggestions and discussions during the preparation of this manuscript. We thank Keats Nelms for his help and use of extracts in the mobility shift assays and Kathleen Conklin and Bob Perry for critical reading of the manuscript.

This work was supported by Public Health Service grant GM37687 from the National Institutes of Health and a grant from the Minnesota Medical Foundation.

LITERATURE CITED

1. **Aebi, M., H. Hormig, R. A. Padgett, J. Reiser, and C. Weissmann.** 1986. Sequence requirements for splicing of higher eucaryotic nuclear pre-mRNA. Cell **47**:555–565.
2. **Alt, F. W., T. K. Blackwell, R. D. DePinho, M. G. Reth, and G. D. Yancopoulos.** 1986. Regulation of genome rearrangement events during lymphocytic differentiation. Immunol. Rev. **89**: 5–30.
3. **Alt, F. W., T. K. Blackwell, and G. D. Yancopoulos.** 1987. Development of the primary antibody repertoire. Science **238**: 1079–1087.
4. **Alwine, J. C., D. J. Kemp, B. A. Parker, J. Reiser, J. Renart, G. R. Stark, and G. M. Wahl.** 1979. Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper. Methods Enzymol. **68**:220–242.
5. **Atchison, M. L., and R. P. Perry.** 1987. The role of the κ enhancer and its binding factor NF-κB in the developmental regulation of κ gene expression. Cell **48**:121–128.
6. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
7. **Blackwell, T. K., and F. W. Alt.** 1989. Molecular characterization of the lymphoid V(D)J recombination activity. J. Biol. Chem. **264**:10327–10330.
8. **Blackwell, T. K., M. W. Moore, G. D. Yancopoulos, H. Suh, S. Lutzker, E. Selsing, and F. W. Alt.** 1986. Recombination between immunoglobuline variable region gene segments is enhanced by transcription. Nature (London) **324**:585–589.
9. **Calame, K. L.** 1985. Mechanism that regulates immunoglobulin gene expression. Annu. Rev. Immunol. **3**:159–195.
10. **Choi, E., M. Kuehl, and R. Wall.** 1980. RNA splicing generates a variant light chain from an aberrantly rearranged κ gene. Nature (London) **286**:776–779.
11. **Chomczynski, P., and N. Sacchi.** 1987. Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. **162**:156–159.
12. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II and a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
13. **Falkner, F. G., and H. G. Zachau.** 1984. Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. Nature (London) **310**: 71–74.
14. **Hromas, R., and B. G. Van Ness.** 1986. Nuclear factors bind to regulatory regions of the mouse kappa immunoglobulin genes. Nucleic Acids Res. **14**:4837–4848.
15. **Jeffreys, A. J., and R. A. Flavell.** 1977. A physical map of the DNA regions flanking the rabbit β-globin genes. Cell **12**:429–439.
16. **Kelley, D. E., and R. P. Perry.** 1986. Transcriptional and posttranscriptional control of immunoglobulin mRNA production during B lymphocyte development. Nucleic Acids Res. **14**:5431–5447.
17. **Kelley, D. W., B. A. Pollock, M. L. Atchison, and R. P. Perry.** 1988. The coupling between enhancer activity and hypomethylation of κ immunoglobulin genes is developmentally regulated. Mol. Cell. Biol. **8**:930–937.
18. **Landolfi, N. F., J. D. Capra, and P. W. Tucker.** 1986. Induction of cell-type-specific nuclear proteins with immunoglobulin V_H promoter region sequence. Nature (London) **323**:548–551.
19. **Landolfi, N. F., J. D. Capra, and P. W. Tucker.** 1987. Protein-

nucleotide contacts in the immunoglobulin heavy-chain promoter region. Proc. Natl. Acad. Sci. USA **84**:3851–3855.
20. **Leclercq, L., P. Butkeraitis, and M. Reth.** 1989. A novel germ-line J_κ transcript starting immediately upstream of $J_\kappa 1$. Nucleic Acids Res. **17**:6809–6819.
21. **Lennon, G. G., and R. P. Perry.** 1985. C_μ-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5′-nontranslatable exon. Nature (London) **318**: 475–478.
22. **Lowery, D. E., and B. G. Van Ness.** 1988. Comparison of in vitro and in vivo splice site selection in κ-immunoglobulin precursor mRNA. Mol. Cell. Biol. **8**:2610–2619.
23. **Lutzker, S., and F. W. Alt.** 1988. Structure and expression of germ line immunoglobulin γ2b Transcripts. Mol. Cell. Biol. **8**:1849–1852.
24. **Martin, D. J., and B. G. Van. Ness.** 1989. Identification of a germ line transcript from the unrearranged kappa gene in human B cells. Mol. Cell. Biol. **9**:4560–4562.
25. **Max, E. E., J. V. Maizel, and P. Leder.** 1981. The nucleotide sequence of a 5.5-kilobase DNA segment containing the mouse κ immunoglobulin J and C region genes. J. Biol. Chem. **256**: 5116–5120.
26. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**: 419–460.
27. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA-binding proteins. Science **245**:371–378.
28. **Nelson, K. J., J. Haimovich, and R. P. Perry.** 1983. Characterization of productive and sterile transcripts from the immunoglobulin heavy-chain locus: processing of μ_m and μ_s mRNA. Mol. Cell. Biol. **3**:1317–1332.
29. **Nelson, K. J., D. E. Kelley, and R. P. Perry.** 1985. Inducible transcription of the unrearranged κ constant-region locus is a common feature of pre-B cells and does not require DNA or protein synthesis. Proc. Natl. Acad. Sci. USA **82**:5305–5309.
30. **Padgett, R. A., P. J. Grabowski, M. M. Konarska, S. Seiler, and P. A. Sharp.** 1986. Splicing of messenger RNA precursors. Annu. Rev. Biochem. **55**:1119–1150.
31. **Parslow, T. G., D. L. Blair, W. J. Murphy, and D. K. Granner.** 1984. Structure of the 5′ ends of immunoglobulin genes: a novel conserved sequence. Proc. Natl. Acad. Sci. USA **74**:560–564.
32. **Parslow, T. G., S. D. Jones, B. Bond, and K. R. Yamamoto.** 1989. The immunoglobulin octanucleotide: independent activity and selective interaction with enhancers. Science **235**:1498–1500.
33. **Perry, R. P., D. E. Kelley, C. Coleclough, J. G. Seidman, P. Leder, S. Tonegawa, G. Matthyssens, and M. Weigert.** 1980. Transcription of mouse κ chain genes: implications for allelic exclusion. Proc. Natl. Acad. Sci. USA **77**:1937–1941.
34. **Reed, R., and T. Maniatis.** 1986. A role for exon sequences and splice-site proximity in splice-site selection. Cell **46**:681–690.
35. **Reth, M. G., and F. W. Alt.** 1984. Novel immunoglobulin heavy chains are produced by DJ_H gene segment rearrangements in lymphoid cells. Nature (London) **312**:418–423.
36. **Reth, M. G., P. Ammirati, S. Jackson, and F. W. Alt.** 1985. Regulated progression of a cultured pre-B cell line to the B-cell stage. Nature (London) **317**:353–355.
37. **Scheidereit, C., A. Heguy, and R. G. Roeder.** 1987. Identification and purification of a human lymphoid-specific octamer binding protein (OTF-2) that activates transcription of an immunoglobulin promoter in vitro. Cell **51**:783–793.
38. **Schlissel, M. S., and D. Baltimore.** 1989. Activation of immunoglobulin kappa gene rearrangement correlates with induction of germline kappa gene transcription. Cell **58**:1001–1007.
39. **Schroeder, H., J. Hillson, and R. Perlmutter.** 1989. Early restriction of the human antibody repertoire. Science **238**:791–793.
40. **Sen, R., and D. Baltimore.** 1987. In vitro transcription of immunoglobulin genes in a B-cell extract: effects of enhancer and promoter sequences. Mol. Cell. Biol. **7**:1989–1994.
41. **Shapiro, M. B., and P. Senapathy.** 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res. **15**:7155–7174.
42. **Shlomchik, M. J., D. A. Nemazee, V. L. Sato, J. VanSnick, D. A. Carson, and M. G. Weigert.** 1986. Variable region sequences of murine IgM anti-IgG monoclonal autoantibodies (rheumatoid factors). J. Exp. Med. **164**:407–427.
43. **Singh, H., R. Sen, D. Baltimore, and P. A. Sharp.** 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature (London) **319**:154–158.
44. **Staudt, L. M., H. Singh, R. Sen, T. Wirth, P. A. Sharp, and D. Baltimore.** 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature (London) **323**: 640–643.
45. **Stavnezer, J., G. Radcliffe, Y.-D. Lin, J. Nietupski, L. Berggren, R. Sitia, and E. Severinson.** 1988. Immunoglobulin heavy-chain switching may be directed by prior induction of transcripts from constant-region genes. Proc. Natl. Acad. Sci. USA **85**:7704–7708.
46. **Tonegawa, S.** 1983. Somatic generation of antibody diversity. Nature (London) **302**:575–581.
47. **Van Ness, B. G., C. Coleclough, R. P. Perry, and M. Weigert.** 1982. DNA between variable and joining gene segments of immunoglobulin κ light chain is frequently retained in cells that rearrange the κ locus. Proc. Natl. Acad. Sci. USA **79**:262–266.
48. **Van Ness, B. G., M. Weigert, C. Coleclough, E. L. Mather, D. E. Kelley, and R. P. Perry.** 1981. Transcription of the unrearranged mouse C_κ locus: sequence of the initiation region and comparison of activity with a rearranged V_κ-C_κ gene. Cell **27**:593–602.
49. **Weaver, D., and D. Baltimore.** 1987. B-lymphocyte-specific protein binding near an immunoglobulin κ-chain gene J segment. Proc. Natl. Acad. Sci. USA **84**:1516–1520.
50. **Wirth, T., L. Staudt, and D. Baltimore.** 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. Nature (London) **329**:174–177.
51. **Wood, D. L., and C. Coleclough.** 1984. Different joining region J elements of the murine κ immunoglobulin light-chain locus are used at markedly different frequencies. Proc. Natl. Acad. Sci. USA **81**:4756–4760.
52. **Yancopoulos, G. D., and F. W. Alt.** 1985. Developmentally controlled and tissue-specific expression of unrearranged V_H gene segments. Cell **40**:271–281.
53. **Yancopoulos, G. D., and F. W. Alt.** 1986. Regulation of the assembly and expression of variable-region genes. Annu. Rev. Immunol. **4**:339–368.
54. **Zhuang, Y., and A. M. Weiner.** 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell **46**:827–835.